APPENDIX C
CRITIQUE OF THE ENGEL & ROTHBARTH METHODOLOGIES

ENGEL — COMMON SENSE OR NONSENSE?

The most widely used method to infer the relative standard of living between households is based upon two observations made by Ernst Engel in 1895. He observed that as a household of a given size became wealthier, the share of the household's total expenditures devoted to food fell. He also observed that as the size of the household increased holding total spending constant, the household was less wealthy and the budget share of food increased. These two observations led him and others to infer that the food budget share was inversely related to the household's economic well being, and that comparisons of households of different sizes and compositions could be made through the comparison of the household's food shares. By finding the level of total spending that equates the food shares across families, the Engel method determines how much more total spending would be needed for a family with N members to be equally well off as a single individual. These proportional factors will be denoted as the Engel equivalence scales.

To formalize Engel's observations, it will be assumed that the log of the food share is a linear function of the log of family size and the log of per capita total spending,

$$ln\left(\frac{f}{X}\right) = a_f + d_f\, ln(N) + b_f\, ln\left(\frac{X}{N}\right) \qquad (1)$$

or

$$ln\left(\frac{f}{X}\right) = a_f + (d_f - b_f)\, ln(N) + b_f\, ln(X) \qquad (1')$$

Engel's first observation that food share declines with X holding family size constant (N), requires that food is a necessity or b is negative. Engel's second observation is that if spending is held constant (X), increases in family size will increase the food share. Hence (d_f-b_f) must be positive. Note that Engel's second observation does not determine the sign of d_f.

Using equation (1), the Engel equivalence scale (M_E) would equal

$$M_E = N^{-\frac{d_f - b_f}{b_f}} = N^{1-\frac{d_f}{b_f}}. \qquad (2)$$

From Engel's two observations, it can be inferred that $-(d_f-b_f)/b_f$ will be positive, but the data will determine whether or not $-(d_f-b_f)/b_f$ is less than, equal to, or greater than one. Common sense would suggest that larger families should enjoy some economies of scale in consumption compared to smaller families, and hence $-(d_f-b_f)/b_f$ should be strictly less than one. This will occur if and only if d_f is less than zero.

At this point, a slight digression will be made about the terminology that will be used. Specifically, the concept of a commodity specific scale will be used to denote the household technology as a function of family size. This scale will equal

$$f_j = N^{1-s_j}$$

where s_j is the scale elasticity of the j[th] commodity. The scale elasticity of a commodity represents how the good can be shared among family members. If s_j is equal to zero, then the good is a private good that cannot be shared and must be replicated if all family members are to enjoy the good to the same degree as a single individual. If s_j is equal to one, then the good is a pure public good that can enjoyed by all members of the family equally. Clearly, most goods will have a scale elasticity that is between zero and one. However, it is possible for diseconomies of scale to occur and this would be represented by a negative value for the scale elasticity.

While commodity specific scale elasticities are interesting, it is most important to know how much more total spending a family of size N requires to be equally well off compared to a single individual, the reference family unit. This will be denoted as the overall equivalence scale and will be equal to

$$f_o = N^{1-s_o}$$

where s_o is the scale elasticity of total spending. If the Engel method is to yield the true overall equivalence scale, then s_o must be equal to d/b or it must be the case that d is equal to bs_o.

Although the exact relationship between the commodity specific and the overall scale elasticities has not been proven, a good candidate would be that the overall scale elasticity is the budget share weighted sum of the commodity specific scales. In other words,

$$s_o = \sum q_k s_k$$

where q_k is the budget share of the k[th] commodity.

As has been demonstrated, the overall scale elasticity of consumption from the Engel method will be estimated by d_f/b_f. While b_f is equal to the income elasticity of the food share and should be negative, what interpretation for d_f can be provided? In a recent paper by Deaton and Paxson (1998), the authors construct a model from which an interpretation can be derived. They begin by assuming there are only two goods — food and non food — and the Barten framework. Specifically, they assume that the family will maximize their well being subject to a total spending constraint. This can be formalized as

$$Maximize \quad U = N\ U\left(\frac{q_f}{f_f}, \frac{q_{nf}}{f_{nf}}\right)$$

$$subject\ to: \quad p_f q_f + p_{nf} q_{nf} = X\ .$$

where

q_i =		the quantity of ith good;
p_i =		the price of the ith good; and
X	=	total spending.

As Deaton and Muellbauer (1980) have shown, the demand for food can be written as

$$q_f = f_f \times g_f\left(f_f p_f, f_{nf} p_{nf}, X\right).$$

where g_f is the food demand for a single individual. Dividing both sides of food demand by N and employing the homogeneity property of the demand, the food demand equation can be rewritten as

$$\frac{q_f}{N} = \frac{f_f}{N} \times g_f\left(\frac{f_f}{N} p_f, \frac{f_{nf}}{N} p_{nf}, \frac{X}{N}\right) \tag{3}$$

Taking logs of both sides of equation (3) and differentiating it with respect to *ln(N)* holding per capita total spending constant, the yield, after some manipulations, is this

$$\left.\frac{\P\, ln(f/N)}{\P\, ln(N)}\right|_{\overline{X/N}} = \left[s_{nf}(1-q_f) + s_f q_f\right] e_{fx} + \left(s_{nf} - s_f\right) s_{ff} - s_f \tag{4}$$

where

s_{ff}	=	the own price substitution elasticity of food ($s_{ff} < 0$);
e_{fx}	=	the income elasticity of food; and
q_f	=	the food budget share.

Equation (4) is similar to Deaton and Paxson's equation (4) but rewritten using the Slutsky relationships between compensated and uncompensated price effects.

Returning to the food share equation (1), if the *ln(X/N)* is added to both sides, the yield is the corresponding per capita food demand

$$ln\left(\frac{f}{N}\right) = a_f + d_f ln(N) + (1 + b_f) ln\left(\frac{X}{N}\right) \qquad (1'')$$

Thus d_f should be estimating equation (4), which can be written as

$$d_f = \left.\frac{\P\, ln(f/N)}{\P\, ln(N)}\right|_{X/N} = s_o(1 + b_f) + \left(s_{nf} - s_f\right) s_{ff} - s_f \qquad (4')$$

Before examining the question of identification of the model, an interpretation of the three components of equation (4') will be provided.

Consider the following hypothetical change: double the size of the family and its total spending. This change will leave the family's per capita spending *(X/N)* unchanged, but increase family size (N). The effect on per capita food expenditures (equation (4')) will be composed of three separate effects. First, if there are any positive scale economies in either food or non food consumption with the increase in family size, then the family will be made better off. With the rise in real income, per capita food consumption should rise. This effect is captured by the term $s_o(1+b_f)$ and is what should be estimated.

However, there are two other effects contained in d_f. In general, it can be assumed that the scale economies of food are not equal to non-food items. For example, housing is one component of non-food that can shared more easily than food. Transportation would most likely fall somewhere between food and housing. To the extent that sharing of non-food items is greater than sharing of food, the relative price of food will rise and the family will substitute away from food, with the result that the per capita food consumption will fall. This effect is captured by the term $(s_{nf}-s_f)\ s_{ff}$. The final effect, σ_f, reflects the direct effect of the scale economies on per capita food consumption and, like the previous effect, will tend to depress per capita food consumption.

Equation (4') highlights the identification problem inherent in the estimation of scale economies in family consumption. The coefficient on *ln(N)* contains four parameters, s_o, s_f,, b and s_{ff} of which only b_f can be determined from the data directly.[1] Thus, if the overall scale elasticity (s_o) is to be determined from data, extra information or assumptions about s_f and s_{ff} will be needed.

[1] Note that s_{nf} is equal to $(s_o-q_f s_f)/(1-q_f)$, hence if s_o and s_f are known, then s_{nf} is also known.

One such set of assumptions was suggested by Gorman (1976) who argued that if the Engel method was to be consistent with the Barten model, all scale elasticities for all commodities must be equal. If the scale elasticity of food is equal to that of non-food items, then there will no price effects of family size and the coefficient on *ln(N)* will equal

$$d_f = s_o(1+b_f) - s_f = b s_o < 0 \ .$$

and be negative. By equating food shares (dividing d_f by b_f), the estimated overall scale elasticity would equal the true elasticity, σ_o.

The appropriateness of the Gorman assumption has always been questioned because few believed that the scale elasticity of food would be the same as the scale elasticity of other goods. Implicitly, most believed that there should be some price effects, although the magnitude of these effects on behavior would be small because s_{ff} was small. Hence, another way to justify the use of food consumption to estimate the overall scale elasticity would be to assume that s_{ff} was zero as in the Prais-Houthakker model of consumer behavior. However, to identify the overall scale elasticity, an additional assumption will have to be made that there are no scale economies in food consumption (s_f is zero). If these two assumptions were true, then

$$d_f = s_o(1+b_f) > 0 \ .$$

Deaton and Paxson proposed these two identifying assumptions as justifying the Engel method. But to identify s_o, one would divide d_f by $(1+b_f)$. In other words, one would equate per capita food consumption across the different family types. Deaton and Paxson argue for per capita food consumption as an indicator of well being and they are not consistent with the Engel method.

Before proceeding to the empirical evidence, let me summarize the two competing sets of identifying assumptions.

	Gorman	Deaton-Paxson
Assumptions	$s_f = s_{nf}$	$s_{ff}=0$ and $s_f=0$
δ	$b s_o < 0$	$(1+b) s_o > 0$
Equate	food shares	per capita food consumption

Deaton and Paxson's estimates for the U.S. show exactly what others have found previously; the estimated coefficient on *ln(N)* which holds *ln(X/N)* constant is negative or close to zero.[2] This empirical finding provides sufficient evidence that the Deaton-Paxson identifying assumptions cannot be appropriate. The dilemma that results is the choice of two bad alternatives. The Gorman assumptions that are most likely inappropriate can be relied upon, or the Engel method can be used with full knowledge that biased estimates of the overall scale elasticity will result.

To judge the sign and the magnitude of the bias (estimate minus truth), the following relationship is used

$$bias = \frac{s_o - s_f}{1 - q_f} \times \left[\frac{1 - q_f + s_{ff}}{b_f} \right].$$

While the first term of the bias will be positive (in general, it would be reasonable to expect that s_o to exceed s_f), the sign of the second term cannot be determined without some idea about the relative magnitude of the substitution elasticity. If the substitution elasticity is less in absolute value than $(1-q_f)$, which is most likely the case, then the bias will be negative — the estimate will be smaller than the true value. If the estimate of the scale elasticity is less than true value, then the amount of compensation implied by Engel's method will overstate the true amount of compensation required to make the families equally well off. This result is consistent with the finding of Deaton and Muellbauer (1986). However, if the substitution elasticity is large, then there will be a positive bias in the estimation of the overall scale elasticity and hence needed compensation would be understated.

ROTHBARTH TO THE RESCUE?

Rothbarth's approach to the estimation of scale economies has a limited scope compared to the ambitions of Engel. By assuming the presence of a good consumed only by the adults in the family, Rothbarth deduced that one could estimate the overall scale economies of consumption (s_o) by examining how parents reduced their consumption of adult goods, q_A. These goods should have the property that additional amounts of the good are not needed when children are present. Thus, if attention is restricted to either one or two parent families (but not together), then increases in family size will represent increases in the number of children. Hence by definition, the commodity specific equivalence scale, f_A, would be 1.0 for all family sizes. This would imply that s_A would be equal to 1.0.

Note a slight difference in the Rothbarth approach in the interpretation of the scale. In the previous interpretation, a scale elasticity of 1.0 implied a pure public good, but in this case, the adult goods are extremely private goods, consumed only by adults and additional amounts are not required when children are present.

[2] The others include Betson (1990), Espenshade (1984) and Watts (1977).

To formalize the Rothbarth approach, it will be assumed that the demand for adult goods has a functional form similar to that of food:

$$ln\left(q_A\right) = a_A + d_A ln(N) + b_A ln\left(\frac{X}{N}\right)$$

Given that adult goods should be a normal good, b_A will be positive. If adults are reducing their consumption of adult goods as the number of children present increases, then $(d_A\text{-}b_A)$ should be negative. Rothbarth conjectured that the level of consumption of adult goods was a good indicator of the adult's well being as well as the well being of the family. If adult consumption is equated, then the Rothbarth equivalence scale would be equal to

$$M_R = N^{-\frac{d_A - b_A}{b_A}} = N^{1-\frac{d_A}{b_A}}.$$

Note again the overall scale elasticity is being estimated by d_A/b_A.

Equations (3) and (4') for adult goods can be written as

$$q_A = g_f\left(\frac{1}{N}p_A, \frac{f_{NA}}{N}p_{NA}, \frac{X}{N}\right)$$

and

$$\left.\frac{\partial\, ln(q_A)}{\partial\, ln(N)}\right|_{\overline{X/N}} = \left[s_{NA}(1-q_A)+q_A\right]b_A + \left(s_{NA}-1\right)s_{AA}$$

$$d_A = s_o b_A + \left(s_{NA}-1\right)s_{AA}$$

where the subscript *NA* denotes non-adult goods. Thus, if the own price substitution elasticity of adult goods (s_{AA}) is zero, then the Rothbarth approach will yield an unbiased estimate of the overall scale elasticity. However, to the extent that adults are price sensitive, then the d_A/b_A will be biased upward $(s_{NA}<1)$ and consequently, the additional amount of total spending that is needed to compensate the adults for their children will be understated.

As in the case of the Engel method, the coefficient on *ln(N)* is predicted to be positive. But, unlike the results from the estimation of the food share equation, my empirical work — Betson (1990) and the current report — estimate the coefficient on *ln(N)* to be positive and less than the income elasticity (b_A). In this case, the empirical findings are consistent with the theory and Rothbarth's conjectures. However, the 'test' of the assumption of zero substitution effects is not much of a test since the substitution effect of additional children reinforces the income effect of the scale economies.

LESSONS

If nonsense is basing ones analysis on a premise that is not true but proceeding anyway, then the Engel method is clearly nonsense. But if nonsense means that no information can be gained from the method, then calling the Engel method nonsense is reaching too far. Clearly the Engel method yields a biased estimate of the overall scale elasticity. And the sign of the bias can reasonably be determined to be negative. Thus, the equivalence scale derived by the Engel method will be overstated. But even a biased estimate of the scale provides some information even if that information is what might be construed as an upper bound for the truth. If it is assumed the welfare being maximized in the Engel representation is identical if not close to the welfare being maximized in the Rothbarth model (which it is not), then what results is the Deaton-Muellbauer lesson that the Engel and Rothbarth estimates provide brackets for the 'truth.'